

Weak Limit of the Geometric Sum of Independent But Not Identically Distributed Random Variables

Alexis Akira Toda*

January 23, 2012

Abstract

We show that when $\{X_j\}$ is a sequence of independent (but not necessarily identically distributed) random variables which satisfies a condition similar to the Lindeberg condition, the properly normalized geometric sum $\sum_{j=1}^{\nu_p} X_j$ (where ν_p is a geometric random variable with mean $1/p$) converges in distribution to a Laplace distribution as $p \rightarrow 0$. The same conclusion holds for the multivariate case. This theorem provides a reason for the ubiquity of the double power law in economic and financial data.

1 Introduction

Let $\{X_j\}$ be a sequence of independent (but not necessarily identically distributed) random variables and ν_p be a geometric random variable with mean $1/p$ independent of X_j 's. The geometric sum

$$\sum_{j=1}^{\nu_p} X_j \quad (1.1)$$

naturally arises in diverse fields [6], particularly in economics. For example, let W_j be the financial wealth of a typical individual at age j and suppose that the wealth grows in a multiplicative way according to $W_{j+1} = G_j W_j$, where G_j is the growth rate which is a random variable. Assuming that each individual dies with constant probability p at each period (and a new individual is born), what does the cross-sectional distribution of wealth look like? To answer this question, let ν_p be a geometric random variable that represents the age of the individual. Letting $X_j = \log G_{j-1}$, the log wealth is

$$\log W_{\nu_p} = \log W_0 + \sum_{j=0}^{\nu_p-1} \log G_j = \log W_0 + \sum_{j=1}^{\nu_p} X_j,$$

*Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511, USA. Email: alexisakira.toda@yale.edu. This paper benefited from comments by Tomasz Kozubowski and Kenichiro Tanaka. The financial supports from the Cowles Foundation, the Nakajima Foundation, and Yale University are greatly acknowledged.

a geometric sum. It might be plausible to assume that $\{X_j\}$ is independent conditional on the realization of macro variables (GDP, interest rate, stock market returns, etc.), but since every individual is more or less affected by the state of the macroeconomy, it is not plausible to assume that $\{X_j\}$ (a time series) is identically distributed conditional on macro variables. In that case the determination of the cross-sectional distribution of log wealth, or the weak limit of the geometric sum (1.1), becomes a non-trivial problem.

The weak limit of the properly normalized random sum (1.1) (where ν_p is not necessarily a geometric random variable but a general integer-valued random variable) has been studied by a number of authors (see [5] and the references therein). In particular, when ν_p is a geometric random variable and X_j has a finite variance, the weak limit of the properly normalized geometric sum (1.1) is a Laplace distribution [10, 11], which has been applied to modeling financial data [13, 9, 8]. However, the literature on the asymptotic distribution of geometric sums seems to be limited to the i.i.d. case. [16] and [5] consider the asymptotic distribution of the random sum of independent but not identically distributed random variables and provide necessary and sufficient conditions for convergence, but since they do not provide explicit examples on geometric sums, it is not obvious whether their general theory applies to the specific case of geometric sums. In this paper by using a technique similar to the proof of the Lindeberg-Feller central limit theorem, I show that the results for the geometric sum of i.i.d. random variables extend to the case when the random variables are independent but not identically distributed (i.n.i.d.).

Before proceeding to the main result we introduce some notations. A random variable X is said to be Laplace if it has a probability density function of the form

$$f(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta}e^{-\alpha|x-m|}, & (x \geq m) \\ \frac{\alpha\beta}{\alpha+\beta}e^{-\beta|x-m|}, & (x < m) \end{cases}$$

where m is the mode and $\alpha, \beta > 0$ are shape parameters. If $\alpha \neq \beta$, X is said to be asymmetric Laplace. The characteristic function of X is

$$\begin{aligned} \phi_X(t) &= \int_{-\infty}^m e^{itx} \frac{\alpha\beta}{\alpha+\beta} e^{-\beta|x-m|} dx + \int_m^{\infty} e^{itx} \frac{\alpha\beta}{\alpha+\beta} e^{-\alpha|x-m|} dx \\ &= \frac{e^{imt}}{1 - i(\frac{1}{\alpha} - \frac{1}{\beta})t + \frac{t^2}{\alpha\beta}}, \end{aligned}$$

from which we obtain the mean $m + \frac{1}{\alpha} - \frac{1}{\beta}$ and the variance $\frac{1}{\alpha^2} + \frac{1}{\beta^2}$. It is often useful to parameterize the Laplace distribution in terms of its characteristic function. Let $a = \frac{1}{\alpha} - \frac{1}{\beta}$ be an asymmetry parameter and $\sigma = \sqrt{\frac{2}{\alpha\beta}}$ be a scale parameter. Then we write $X \sim \mathcal{AL}(m, a, \sigma)$ if

$$\phi_X(t) = \frac{e^{imt}}{1 - iat + \frac{\sigma^2 t^2}{2}}.$$

The mean, mode, and variance of $\mathcal{AL}(m, a, \sigma)$ is $m + a$, m , and $a^2 + \sigma^2$, respectively. In particular, setting $\alpha = \beta = \frac{\sqrt{2}}{\sigma}$, the symmetric Laplace distribution with mean and mode m and standard deviation σ (which we denote by $\mathcal{L}(m, \sigma)$) has density $f(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}|x-m|}{\sigma}}$ and characteristic function $\frac{e^{imt}}{1 + \frac{\sigma^2 t^2}{2}}$. A comprehensive review of the Laplace distribution can be found in [7].

2 Main result

Theorem 2.1. *Let $\{X_j\}$ be a sequence of independent but not identically distributed (i.n.i.d) random variables such that $E[X_j] = 0$ and $\text{Var}[X_j] = \sigma_j^2$, $\{a_j\}$ be a real sequence, and ν_p be a geometric random variable independent of X_j 's with mean $1/p$. Suppose that*

1. $\lim_{n \rightarrow \infty} n^{-\alpha} \sigma_n^2 = 0$ for some $0 < \alpha < 1$ and $\sigma^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \sigma_j^2 > 0$ exists,
2. $a := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n a_j$ exists, and
3. for all $\epsilon > 0$ we have

$$\lim_{p \rightarrow 0} \sum_{j=1}^{\infty} (1-p)^{j-1} p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] = 0. \quad (2.1)$$

Then, as $p \rightarrow 0$ the geometric sum $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)$ converges in distribution to $\mathcal{AL}(0, a, \sigma)$.

By strengthening the assumptions of Theorem 2.1, we obtain the following corollaries.

Corollary 2.2. *Let $\{X_j\}$ be a sequence of independent but not identically distributed (i.n.i.d) random variables such that $E[X_j] = 0$, $\text{Var}[X_j] = \sigma_j^2$, and $\sigma^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \sigma_j^2 > 0$ exists. Let $\{a_j\}$ be a real sequence such that $a := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n a_j$ exists, and ν_p be a geometric random variable independent of X_j 's with mean $1/p$. Suppose that $\{X_j^2\}$ is uniformly integrable. Then $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j) \xrightarrow{d} \mathcal{AL}(0, a, \sigma)$ as $p \rightarrow 0$.*

Proof. For $c > 0$ let $M(c) = \sup_j E[X_j^2 \{ |X_j| \geq c \}]$. Since $\{X_j^2\}$ is uniformly integrable, we have $M(c) \rightarrow 0$ as $c \rightarrow \infty$, so $M(c) < \infty$ for sufficiently large c . For such c , we have

$$\sigma_j^2 = E[X_j^2 \{ |X_j| < c \}] + E[X_j^2 \{ |X_j| \geq c \}] \leq c^2 + M(c),$$

so $\{\sigma_j\}$ is bounded, in particular $n^{-\alpha} \sigma_n^2 \rightarrow 0$ for any $0 < \alpha < 1$. For any $\epsilon > 0$ and $c > 0$ choose p such that $\epsilon p^{-\frac{1}{2}} \geq c$. Then

$$\begin{aligned} \sum_{j=1}^{\infty} (1-p)^{j-1} p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] &\leq \sum_{j=1}^{\infty} (1-p)^{j-1} p E \left[X_j^2 \{ |X_j| \geq c \} \right] \\ &\leq M(c), \end{aligned}$$

so letting $p \rightarrow 0$ and then $c \rightarrow \infty$, condition (2.1) holds. \square

Corollary 2.3. *Let $\{X_j\}$ be a sequence of i.i.d. random variables with mean 0 and variance σ^2 , $\{a_j\}$ a real sequence such that $a := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n a_j$ exists, and ν_p a geometric random variable independent of X_j 's with mean $1/p$. Then $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j) \xrightarrow{d} \mathcal{AL}(0, a, \sigma)$ as $p \rightarrow 0$.*

Proof. Since X_j 's are i.i.d., $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \sigma_j^2 = \sigma^2 > 0$ and $\{X_j^2\}$ is uniformly integrable. Hence the conclusion holds by Corollary 2.2. \square

The proof of Theorem 2.1 is based on the idea of Lindeberg [12] for proving the central limit theorem. We first prove Theorem 2.1 when X_j 's are Gaussian. Then we take a sequence of independent zero mean Gaussian variables $\{Y_j\}$ with the same variances as $\{X_j\}$ and show that the geometric sums $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)$ and $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)$ admit the same weak limit.

Proposition 2.4. *Let $\{Y_j\}$ be a sequence of independent Gaussian random variables such that $E[Y_j] = 0$, $\text{Var}[Y_j] = \sigma_j^2$, and $\sigma^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \sigma_j^2 > 0$ exists. Let $\{a_j\}$ be a real sequence such that $a := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n a_j$ exists, and ν_p be a geometric random variable independent of Y_j 's with mean $1/p$. Then $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j) \xrightarrow{d} \mathcal{AL}(0, a, \sigma)$ as $p \rightarrow 0$.*

Proof. Let $S_n = \sum_{j=1}^n Y_j$, $b_n = \sum_{j=1}^n a_j$, and $\tau_n^2 = \sum_{j=1}^n \sigma_j^2$. Since Y_j 's are independent Gaussian, $S_n \sim N(0, \tau_n^2)$. By conditioning on ν_p the characteristic function of the geometric sum $Z_p := p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)$ is

$$\begin{aligned} \phi_p(t) &:= E[e^{itZ_p}] = E[E[e^{itZ_p} | \nu_p]] = \sum_{n=1}^{\infty} (1-p)^{n-1} p E[e^{itp^{\frac{1}{2}}(S_n + p^{\frac{1}{2}} b_n)}] \\ &= \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{itpb_n - \frac{pt^2 \tau_n^2}{2}} = \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn(-it\frac{b_n}{n} + \frac{t^2 \tau_n^2}{2n})}. \end{aligned}$$

Let $z_n = -it\frac{b_n}{n} + \frac{t^2 \tau_n^2}{2n}$. By assumption, $z_n \rightarrow z := -iat + \frac{\sigma^2 t^2}{2}$ as $n \rightarrow \infty$. Since $\text{Re } z = \frac{\sigma^2 t^2}{2} \geq 0 > -1$, by Lemma A.3, we obtain

$$\lim_{p \rightarrow 0} \phi_p(t) = \frac{1}{1+z} = \frac{1}{1 - iat + \frac{\sigma^2 t^2}{2}}.$$

Hence $Z_p \xrightarrow{d} \mathcal{AL}(0, a, \sigma)$ as $p \rightarrow 0$. \square

Next we show that condition (2.1) holds for $\{Y_j\}$.

Lemma 2.5. *Let everything be as in Theorem 2.1 and $\{Y_j\}$ be as in Proposition 2.4. Then condition (2.1) holds for $\{Y_j\}$.*

Proof. Since $n^{-\alpha} \sigma_n^2 \rightarrow 0$ for some $0 < \alpha < 1$, for any $\delta > 0$ we can choose N such that $n^{-\alpha} \sigma_n^2 \leq \delta$ for $n > N$. Since by assumption $Y_j \sim N(0, \sigma_j^2)$, we have $Z = Y_j / \sigma_j \sim N(0, 1)$. Let $\eta = \frac{2\alpha}{1-\alpha} > 0$ and $c = \epsilon p^{-\frac{1}{2}}$. Since $|Y_j/c| \geq 1$ when $|Y_j| \geq c$, for $j > N$ we obtain

$$\begin{aligned} E[Y_j^2 \{ |Y_j| \geq c \}] &\leq E[Y_j^2 |Y_j/c|^\eta \{ |Y_j| \geq c \}] \\ &\leq E[Y_j^2 |Y_j/c|^\eta] \leq \frac{\sigma_j^{2+\eta}}{c^\eta} E[|Z|^{2+\eta}] \\ &\leq \frac{\delta^{\frac{2+\eta}{2}} j^{\frac{\alpha(2+\eta)}{2}}}{c^\eta} E[|Z|^{2+\eta}] = \frac{\delta^{\frac{1}{1-\alpha}} j^{\frac{\alpha}{1-\alpha}}}{c^\eta} E[|Z|^{2+\eta}], \quad (2.2) \end{aligned}$$

where we have used $\frac{2+\eta}{2} = \frac{1}{1-\alpha}$ by the definition of η . Substituting $c = \epsilon p^{-\frac{1}{2}}$, multiplying (2.2) by $(1-p)^{j-1}p$ and summing over $j > N$, it follows from Lemma A.4 that

$$\begin{aligned} & \sum_{j=N+1}^{\infty} (1-p)^{j-1}p \mathbb{E} \left[Y_j^2 \left\{ |Y_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ & \leq \frac{p}{1-p} \sum_{j=N+1}^{\infty} (1-p)^j \frac{p^{\frac{\eta}{2}} \delta^{\frac{1}{1-\alpha}} j^{\frac{\alpha}{1-\alpha}}}{\epsilon^{\eta}} \mathbb{E}[|Z|^{2+\eta}] \\ & \leq \frac{\delta^{\frac{1}{1-\alpha}} p^{\frac{1}{1-\alpha}}}{\epsilon^{\eta}(1-p)} \mathbb{E}[|Z|^{2+\eta}] \sum_{j=1}^{\infty} (1-p)^j j^{\frac{\alpha}{1-\alpha}} \leq \frac{C \delta^{\frac{1}{1-\alpha}}}{\epsilon^{\eta}(1-p)} \mathbb{E}[|Z|^{2+\eta}] \end{aligned}$$

for some constant $C > 0$. Hence

$$\begin{aligned} & \sum_{j=1}^{\infty} (1-p)^{j-1}p \mathbb{E} \left[Y_j^2 \left\{ |Y_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ & \leq \sum_{j=1}^N (1-p)^{j-1}p \mathbb{E} \left[Y_j^2 \left\{ |Y_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] + \frac{C \delta^{\frac{1}{1-\alpha}}}{\epsilon^{\eta}(1-p)} \mathbb{E}[|Z|^{2+\eta}]. \end{aligned}$$

Letting $p \rightarrow 0$ and then $\delta \rightarrow 0$, condition (2.1) holds for $\{Y_j\}$. \square

Proposition 2.6. *Let $\{X_j\}$, $\{Y_j\}$, $\{a_j\}$, and ν_p be as in Theorem 2.1 and Proposition 2.4. Then for any bounded C^3 function f on \mathbb{R} with bounded derivatives up to the third order, we have*

$$\lim_{p \rightarrow 0} \left| \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)) \right] - \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)) \right] \right| = 0.$$

Proof. Fix n and consider

$$\begin{aligned} & f(p^{\frac{1}{2}} \sum_{j=1}^n (X_j + p^{\frac{1}{2}} a_j)) - f(p^{\frac{1}{2}} \sum_{j=1}^n (Y_j + p^{\frac{1}{2}} a_j)) \\ & = \sum_{j=1}^n [f(p^{\frac{1}{2}} (X_j + Z_j)) - f(p^{\frac{1}{2}} (Y_j + Z_j))], \end{aligned} \quad (2.3)$$

where

$$Z_j = X_1 + \cdots + X_{j-1} + Y_{j+1} + \cdots + Y_n + p^{\frac{1}{2}} \sum_{k=1}^n a_k.$$

By Corollary A.6, the j -th term of (2.3) is equal to

$$\begin{aligned} & f(p^{\frac{1}{2}} (X_j + Z_j)) - f(p^{\frac{1}{2}} (Y_j + Z_j)) \\ & = f'(p^{\frac{1}{2}} Z_j) p^{\frac{1}{2}} (X_j - Y_j) + \frac{f''(p^{\frac{1}{2}} Z_j)}{2} p (X_j^2 - Y_j^2) + R_j, \end{aligned} \quad (2.4)$$

where R_j is the remainder term. R_j is bounded by

$$|R_j| \leq g(p^{\frac{1}{2}} X_j) + g(p^{\frac{1}{2}} Y_j), \quad (2.5)$$

where $g(h) = K \min \{h^2, |h|^3\}$ for some $K > 0$. Noting that X_j, Y_j are independent of Z_j , $E[X_j] = E[Y_j] = 0$, and $\text{Var}[X_j] = \text{Var}[Y_j] = \sigma_j^2$, taking expectations of both sides of (2.4), we get

$$\left| E[f(p^{\frac{1}{2}}(X_j + Z_j))] - E[f(p^{\frac{1}{2}}(Y_j + Z_j))] \right| \leq E[|R_j|].$$

Therefore by the triangle inequality and (2.5) we obtain

$$\begin{aligned} & \left| E[f(p^{\frac{1}{2}} \sum_{j=1}^n (X_j + p^{\frac{1}{2}} a_j))] - E[f(p^{\frac{1}{2}} \sum_{j=1}^n (Y_j + p^{\frac{1}{2}} a_j))] \right| \\ & \leq \sum_{j=1}^n E[|R_j|] \leq \sum_{j=1}^n \left(E[g(p^{\frac{1}{2}} X_j)] + E[g(p^{\frac{1}{2}} Y_j)] \right). \end{aligned} \quad (2.6)$$

Using the definition of g , we can bound $E[g(p^{\frac{1}{2}} X_j)]$ as

$$\begin{aligned} E[g(p^{\frac{1}{2}} X_j)] &= E \left[g(p^{\frac{1}{2}} X_j) \left\{ |X_j| < \epsilon p^{-\frac{1}{2}} \right\} \right] + E \left[g(p^{\frac{1}{2}} X_j) \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ &\leq K E \left[\left| p^{\frac{1}{2}} X_j \right|^3 \left\{ |X_j| < \epsilon p^{-\frac{1}{2}} \right\} \right] + K E \left[(p^{\frac{1}{2}} X_j)^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ &= K E \left[(p^{\frac{1}{2}} X_j)^2 \epsilon \left\{ |X_j| < \epsilon p^{-\frac{1}{2}} \right\} \right] + K E \left[(p^{\frac{1}{2}} X_j)^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ &\leq \epsilon K p \sigma_j^2 + K p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right]. \end{aligned} \quad (2.7)$$

Now let $\tau_n^2 = \sum_{j=1}^n \sigma_j^2$. Since $\tau_n^2/n \rightarrow \sigma^2$ by assumption, $\{\tau_n^2/n\}$ is bounded by some $M > 0$. Then

$$\sum_{n=1}^{\infty} (1-p)^{n-1} p \sum_{j=1}^n \epsilon K p \sigma_j^2 \leq \sum_{n=1}^{\infty} (1-p)^{n-1} p \epsilon K p M n = \epsilon K M. \quad (2.8)$$

Also, by changing the order of summation we obtain

$$\begin{aligned} & \sum_{n=1}^{\infty} (1-p)^{n-1} p \sum_{j=1}^n K p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ &= \sum_{j=1}^{\infty} \sum_{n=j}^{\infty} (1-p)^{n-1} p K p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ &= K \sum_{j=1}^{\infty} (1-p)^{j-1} p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right]. \end{aligned} \quad (2.9)$$

Combining (2.7), (2.8), and (2.9), we obtain

$$\begin{aligned} & \sum_{n=1}^{\infty} (1-p)^{n-1} p \sum_{j=1}^n E[g(p^{\frac{1}{2}} X_j)] \\ & \leq K M \epsilon + K \sum_{j=1}^{\infty} (1-p)^{j-1} p E \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right]. \end{aligned} \quad (2.10)$$

The same inequality as (2.10) holds when $\{X_j\}$ is replaced by $\{Y_j\}$. Hence applying condition (2.1) to (2.10) and invoking Lemma 2.5, it follows from (2.6) and (2.10) that

$$\begin{aligned}
& \left| \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)) \right] - \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)) \right] \right| \\
& \leq \sum_{n=1}^{\infty} (1-p)^{n-1} p \left| \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^n (X_j + p^{\frac{1}{2}} a_j)) \right] - \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^n (Y_j + p^{\frac{1}{2}} a_j)) \right] \right| \\
& \leq \sum_{n=1}^{\infty} (1-p)^{n-1} p \sum_{j=1}^n \left(\mathbb{E}[g(p^{\frac{1}{2}} X_j)] + \mathbb{E}[g(p^{\frac{1}{2}} Y_j)] \right) \\
& \leq 2KM\epsilon + K \sum_{j=1}^{\infty} (1-p)^{j-1} p \left(\mathbb{E} \left[X_j^2 \left\{ |X_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] + \mathbb{E} \left[Y_j^2 \left\{ |Y_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \right) \\
& \rightarrow 2KM\epsilon
\end{aligned}$$

as $p \rightarrow 0$. Since $\epsilon > 0$ is arbitrary, letting $\epsilon \rightarrow 0$ we get

$$\lim_{p \rightarrow 0} \left| \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)) \right] - \mathbb{E} \left[f(p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)) \right] \right| = 0.$$

□

Proof of Theorem 2.1. Let $f(x) = e^{itx}$. f is C^∞ and all of its derivatives are bounded because $|f^{(n)}(x)| = |(it)^n e^{itx}| = t^n$, which does not depend on x . Let $\{Y_j\}$ be as in Proposition 2.4. Then by Proposition 2.6, we get

$$\lim_{p \rightarrow 0} \left| \mathbb{E} \left[e^{itp^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)} \right] - \mathbb{E} \left[e^{itp^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)} \right] \right| = 0.$$

Hence by Proposition 2.4 we have

$$\lim_{p \rightarrow 0} \mathbb{E} \left[e^{itp^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)} \right] = \lim_{p \rightarrow 0} \mathbb{E} \left[e^{itp^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (Y_j + p^{\frac{1}{2}} a_j)} \right] = \frac{1}{1 - iat + \frac{\sigma^2 t^2}{2}}.$$

Since the right-most expression is the characteristic function of $\mathcal{AL}(0, a, \sigma)$ which is continuous at $t = 0$, by Lévy's continuity theorem $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (X_j + p^{\frac{1}{2}} a_j)$ converges in distribution to $\mathcal{AL}(0, a, \sigma)$ as $p \rightarrow 0$. □

3 Multivariate case

The generalization of Theorem 2.1 to the multivariate case is straightforward. If \mathbf{X} is a d -dimensional random variable with characteristic function

$$\phi_{\mathbf{X}}(\mathbf{t}) = \frac{e^{i\mathbf{m}'\mathbf{t}}}{1 - i\mathbf{a}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\mathbf{\Sigma}\mathbf{t}},$$

where $\mathbf{m}, \mathbf{a} \in \mathbb{R}^d$ and $\mathbf{\Sigma}$ is a $d \times d$ symmetric and positive definite matrix, then the distribution of \mathbf{X} is said to be multivariate Laplace which we denote by $\mathcal{AL}_d(\mathbf{m}, \mathbf{a}, \mathbf{\Sigma})$. The mean, mode, and variance of $\mathcal{AL}_d(\mathbf{m}, \mathbf{a}, \mathbf{\Sigma})$ is $\mathbf{m} + \mathbf{a}$, \mathbf{m} , and $\mathbf{\Sigma} + \mathbf{a}\mathbf{a}'$, respectively.

Theorem 3.1. Let $\{\mathbf{X}_j\}$ be a sequence of independent but not identically distributed (i.n.i.d) random vectors in \mathbb{R}^d such that $\mathbb{E}[\mathbf{X}_j] = \mathbf{0}$ and $\text{Var}[\mathbf{X}_j] = \Sigma_j$, $\{\mathbf{a}_j\}$ be a sequence in \mathbb{R}^d , and ν_p be a geometric random variable independent of \mathbf{X}_j 's with mean $1/p$. Suppose that

1. $\lim_{n \rightarrow \infty} n^{-\alpha} \Sigma_n = \mathbf{O}$ for some $0 < \alpha < 1$ and $\Sigma := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \Sigma_j$ exists and positive definite,
2. $\mathbf{a} := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{a}_j$ exists, and
3. for all $\epsilon > 0$ we have

$$\lim_{p \rightarrow 0} \sum_{j=1}^{\infty} (1-p)^{j-1} p \mathbb{E} \left[\|\mathbf{X}_j\|^2 \left\{ \|\mathbf{X}_j\| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] = 0, \quad (3.1)$$

where $\|\cdot\|$ denotes the Euclidean norm.

Then, as $p \rightarrow 0$ the geometric sum $p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (\mathbf{X}_j + p^{\frac{1}{2}} \mathbf{a}_j)$ converges in distribution to $\mathcal{AL}_d(\mathbf{0}, \mathbf{a}, \Sigma)$.

Proof. Let us first show that for any $\mathbf{0} \neq \mathbf{t} \in \mathbb{R}^d$ the sequence of real random variables $\{\mathbf{t}' \mathbf{X}_j\}$ satisfies the assumptions of Theorem 2.1. Since $\mathbb{E}[\mathbf{X}_j] = \mathbf{0}$ and $\text{Var}[\mathbf{X}_j] = \Sigma_j$, we have $\mathbb{E}[\mathbf{t}' \mathbf{X}_j] = 0$ and $\text{Var}[\mathbf{t}' \mathbf{X}_j] = \mathbf{t}' \Sigma_j \mathbf{t}$. Hence $\lim_{n \rightarrow \infty} n^{-\alpha} \text{Var}[\mathbf{t}' \mathbf{X}_n] = 0$ and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \text{Var}[\mathbf{t}' \mathbf{X}_j] = \mathbf{t}' \Sigma \mathbf{t} > 0$$

because Σ is positive definite and $\mathbf{t} \neq \mathbf{0}$. Also, $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{t}' \mathbf{a}_j = \mathbf{t}' \mathbf{a}$. By the Cauchy-Schwarz inequality, we have $|\mathbf{t}' \mathbf{X}| \leq \|\mathbf{t}\| \|\mathbf{X}\|$. Hence

$$\{|\mathbf{t}' \mathbf{X}| \geq c\} \subset \{\|\mathbf{t}\| \|\mathbf{X}\| \geq c\} = \left\{ \|\mathbf{X}\| \geq \frac{c}{\|\mathbf{t}\|} \right\}.$$

Therefore for all $\epsilon > 0$ we have

$$\begin{aligned} & \sum_{j=1}^{\infty} (1-p)^{j-1} p \mathbb{E} \left[(\mathbf{t}' \mathbf{X}_j)^2 \left\{ |\mathbf{t}' \mathbf{X}_j| \geq \epsilon p^{-\frac{1}{2}} \right\} \right] \\ & \leq \|\mathbf{t}\|^2 \sum_{j=1}^{\infty} (1-p)^{j-1} p \mathbb{E} \left[\|\mathbf{X}_j\|^2 \left\{ \|\mathbf{X}_j\| \geq \frac{\epsilon}{\|\mathbf{t}\|} p^{-\frac{1}{2}} \right\} \right] \rightarrow 0 \end{aligned}$$

as $p \rightarrow 0$ by condition (3.1), so $\{\mathbf{t}' \mathbf{X}_j\}$ satisfies condition (2.1) of Theorem 2.1. Since $\{\mathbf{t}' \mathbf{X}_j\}$ satisfies all assumptions of Theorem 2.1, it follows that

$$\mathbf{t}' p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (\mathbf{X}_j + p^{\frac{1}{2}} \mathbf{a}_j) = p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (\mathbf{t}' \mathbf{X}_j + p^{\frac{1}{2}} \mathbf{t}' \mathbf{a}_j) \xrightarrow{d} \mathcal{AL}(\mathbf{0}, \mathbf{t}' \mathbf{a}, \sqrt{\mathbf{t}' \Sigma \mathbf{t}})$$

as $p \rightarrow 0$. This shows that

$$\lim_{p \rightarrow 0} \mathbb{E} \left[e^{i \mathbf{t}' p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (\mathbf{X}_j + p^{\frac{1}{2}} \mathbf{a}_j)} \right] = \frac{1}{1 - i \mathbf{t}' \mathbf{a} + \frac{1}{2} \mathbf{t}' \Sigma \mathbf{t}}. \quad (3.2)$$

In proving (3.2) we have assumed that $\mathbf{t} \neq \mathbf{0}$, but (3.2) trivially holds for $\mathbf{t} = \mathbf{0}$. Since the right-most expression of (3.2) is continuous at $\mathbf{t} = \mathbf{0}$, by Lévy's continuity theorem

$$p^{\frac{1}{2}} \sum_{j=1}^{\nu_p} (\mathbf{X}_j + p^{\frac{1}{2}} \mathbf{a}_j) \xrightarrow{d} \mathcal{AL}_d(\mathbf{0}, \mathbf{a}, \Sigma).$$

□

4 Concluding remarks

In this paper I showed that the properly normalized geometric sum $\sum_{j=1}^{\nu_p} X_j$ converges in distribution to a Laplace random variable even if the random variables $\{X_j\}$ are not identically distributed as long as they are independent. The proof is similar to that of the Lindeberg-Feller central limit theorem. This theorem provides a reason why many economic and financial variables obey the power law not just in the right tail [2] but also in the left tail. If an economic variable results from a large, deterministic number of independent multiplicative shocks, that variable will be lognormally distributed as first observed by [3]. However, in reality many variables seem to be well-described by the double Pareto and related distribution [14, 15, 4, 17, 1]. If we incorporate the death probability of economic units in the model, the number of multiplicative shocks is not deterministic but a geometric random variable. Theorem 2.1 (in exponential form) then states that the geometric product of independent positive random variables tends to the double Pareto distribution, which is empirically supported.

Since the central limit theorem holds under general conditions (for example, ergodicity and stationarity), we can expect that the properly normalized geometric sum of random variables converges in distribution to a Laplace distribution under such conditions even if independence fails. Addressing these issues are beyond the scope of this paper but interesting to pursue.

A Lemmas

Lemma A.1. *For $z \in \mathbb{C}$, we have $|e^z - 1| \leq |z|e^{|z|}$.*

Proof. Using the Taylor expansion of e^z , we obtain

$$|e^z - 1| = \left| \sum_{n=1}^{\infty} \frac{z^n}{n!} \right| \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{(n+1)!} \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{n!} = |z|e^{|z|}.$$

□

Lemma A.2. *For $0 < p < 1$ and $x \geq -1$, we have $0 < (1-p)e^{-px} < 1$.*

Proof. Since $e^t \geq 1 + t$ for all t , we get $e^{px} \geq 1 + px \geq 1 - p > 0$. The first equality holds if and only if $px = 0$ and the second if and only if $x = -1$, but since $0 < p < 1$ the two equalities cannot hold simultaneously. Hence $0 < (1-p)e^{-px} < 1$. □

Lemma A.3. Let $0 < p < 1$ and $\{z_n\} \subset \mathbb{C}$ be such that $\lim_{n \rightarrow \infty} z_n = z$ with $\operatorname{Re} z > -1$. Then

$$\lim_{p \rightarrow 0} \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn z_n} = \frac{1}{1+z}.$$

Proof. For $0 < p < 1$ let $S(p) = \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn z_n}$. First we prove that $S(p)$ exists. For this purpose let $z_n = x_n + iy_n$. Since $\lim x_n > -1$, we can choose $N > 0$ such that $x_n > -1$ for $n > N$. Then by the triangle inequality

$$\begin{aligned} |S(p)| &\leq \sum_{n=1}^N (1-p)^{n-1} p e^{-pn x_n} + \sum_{n=N+1}^{\infty} (1-p)^{n-1} p e^{-pn x_n} \\ &\leq \sum_{n=1}^N (1-p)^{n-1} p e^{-pn x_n} + \sum_{n=N+1}^{\infty} (1-p)^{n-1} p e^{pn} \\ &= \sum_{n=1}^N (1-p)^{n-1} p e^{-pn x_n} + (1-p)^N \frac{p e^p}{1 - (1-p)e^p} \end{aligned}$$

because $0 < (1-p)e^p < 1$ by setting $x = -1$ in Lemma A.2. Hence $S(p)$ exists. Replacing $\{z_n\}$ with z and applying the same argument,

$$T(p) := \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn z} = \frac{p e^{-pz}}{1 - (1-p)e^{-pz}}$$

exists. Now by l'Hôpital's rule we have

$$\lim_{p \rightarrow 0} T(p) = \lim_{p \rightarrow 0} \frac{p}{e^{pz} - (1-p)} = \lim_{p \rightarrow 0} \frac{1}{z e^{pz} + 1} = \frac{1}{1+z},$$

so it suffices to show that $|S(p) - T(p)| \rightarrow 0$ as $p \rightarrow 0$. For any $0 < \epsilon < 1 + \operatorname{Re} z$, choose $N > 0$ such that $|z_n - z| < \epsilon$ for $n > N$. Consider

$$\begin{aligned} |S(p) - T(p)| &\leq \sum_{n=1}^N (1-p)^{n-1} p |e^{-pn z_n} - e^{-pn z}| \\ &\quad + \sum_{n=N+1}^{\infty} (1-p)^{n-1} p |e^{-pn z_n} - e^{-pn z}| = I + II. \end{aligned}$$

Since each term of $S(p)$ and $T(p)$ tends to zero as $p \rightarrow 0$, we have $I \rightarrow 0$. By the choice of N and Lemma A.1, letting $z = x + iy$ we get

$$\begin{aligned} II &\leq \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn x} \left| e^{-pn(z_n - z)} - 1 \right| \leq \sum_{n=1}^{\infty} (1-p)^{n-1} p e^{-pn x} p n |z_n - z| e^{pn |z_n - z|} \\ &\leq \sum_{n=N+1}^{\infty} (1-p)^{n-1} p e^{-pn x} p n \epsilon e^{pn \epsilon} = \sum_{n=N+1}^{\infty} (1-p)^{n-1} p e^{-pn(x-\epsilon)} p n \epsilon. \end{aligned}$$

Since $\operatorname{Re} z - \epsilon > -1$, by Lemma A.2 we have $(1-p)e^{-p(x-\epsilon)} < 1$, so the above

sum converges. Then

$$\begin{aligned} II &\leq \epsilon p^2 \sum_{n=N+1}^{\infty} (1-p)^{n-1} e^{-pn(x-\epsilon)} n \\ &= \epsilon \left(\frac{p}{e^{p(x-\epsilon)} - 1 + p} \right)^2 (1-p)^N e^{-p(N-1)(x-\epsilon)} \rightarrow \frac{\epsilon}{(1+x-\epsilon)^2} \end{aligned}$$

as $p \rightarrow 0$ by applying l'Hôpital's rule to the above fraction. Letting $\epsilon \rightarrow 0$, we obtain $II \rightarrow 0$. Hence $\lim S(p) = \lim T(p) = \frac{1}{1+z}$. \square

Lemma A.4. *If $0 < p < 1$ and $\alpha > -1$, there exists a constant $C > 0$ such that*

$$\sum_{n=1}^{\infty} (1-p)^n n^\alpha \leq Cp^{-\alpha-1}.$$

Proof. Let $f(x) = (1-p)^x x^\alpha$. If $-1 < \alpha \leq 0$, $f(x)$ is monotone decreasing for $x \geq 0$. Since $\log(1-p) \leq -p$, we obtain

$$\sum_{n=1}^{\infty} (1-p)^n n^\alpha \leq \int_0^{\infty} (1-p)^x x^\alpha dx = \frac{\Gamma(\alpha+1)}{(-\log(1-p))^{\alpha+1}} \leq Cp^{-\alpha-1},$$

where $C = \Gamma(\alpha+1)$. If $\alpha > 0$, again using $\log(1-p) \leq -p$ we have

$$f'(x) = [\alpha + \log(1-p)x](1-p)^x x^{\alpha-1} \leq 0$$

for $x \geq \frac{\alpha}{p}$. Hence we obtain

$$\begin{aligned} \sum_{n=1}^{\infty} (1-p)^n n^\alpha &= \sum_{n \leq \alpha/p} (1-p)^n n^\alpha + \sum_{n > \alpha/p} (1-p)^n n^\alpha \\ &\leq \left(\frac{\alpha}{p}\right) \left(\frac{\alpha}{p}\right)^\alpha + \int_0^{\infty} (1-p)^x x^\alpha dx \\ &= \left(\frac{\alpha}{p}\right)^{\alpha+1} + \frac{\Gamma(\alpha+1)}{(-\log(1-p))^{\alpha+1}} \leq Cp^{-\alpha-1}, \end{aligned}$$

where $C = \alpha^{\alpha+1} + \Gamma(\alpha+1)$. \square

Lemma A.5. *Let f be a bounded C^3 function on \mathbb{R} with bounded derivatives up to the third order. Then there exists a constant K such that for all $h \in \mathbb{R}$, we have*

$$g(h) := \sup_{x \in \mathbb{R}} \left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right| \leq K \min \{h^2, |h|^3\}.$$

Proof. By assumption $M_i := \sup_{x \in \mathbb{R}} |f^{(i)}(x)| < \infty$ for $i = 0, 1, 2, 3$. By Taylor's theorem for each $x, x+h \in \mathbb{R}$ there exists ξ between x and $x+h$ such that

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{f'''(\xi)}{6}h^3.$$

Hence

$$\left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right| \leq \frac{M_3}{6} |h|^3 =: K_1 |h|^3.$$

On the other hand, by the triangle inequality we get

$$\left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right| \leq 2M_0 + M_1|h| + \frac{M_2}{2}|h|^2 \leq K_2|h|^2$$

for large enough $|h|$, say $|h| \geq b$. Then for $|h| < b$ we have

$$\left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right| \leq K_1|h|^3 \leq K_1b|h|^2.$$

Hence by taking $K = \max\{K_1, K_2, K_1b\}$ we obtain

$$\left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right| \leq K \min\{h^2, |h|^3\}.$$

□

Corollary A.6. *Let everything be as in Lemma A.5. Then*

$$f(x+h_1) - f(x+h_2) = f'(x)(h_1 - h_2) + \frac{f''(x)}{2}(h_1^2 - h_2^2) + R(x, h_1, h_2),$$

where the remainder term R satisfies

$$|R(x, h_1, h_2)| \leq g(h_1) + g(h_2) \leq K \left[\min\{h_1^2, |h_1|^3\} + \min\{h_2^2, |h_2|^3\} \right]$$

Proof. Trivial by Lemma A.5 and the triangle inequality. □

References

- [1] Simone Alfarano, Mishaël Milaković, Albrecht Irle, and Jonas Kauschke. A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149, 2012.
- [2] Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–293, 2009.
- [3] Robert Gibrat. *Les Inégalités économiques*. Librairie du Recueil Sirey, Paris, 1931.
- [4] Kristian Giesen, Arndt Zimmermann, and Jens Suedekum. The size distribution across all cities—double Pareto lognormal strikes. *Journal of Urban Economics*, 68:129–137, 2010.
- [5] Boris V. Gnedenko and Victor Yu. Korolev. *Random Summation: Limit Theorems and Applications*. CRC Press, Boca Raton, FL, 1996.
- [6] Vladimir V. Kalashnikov. *Geometric Sums: Bounds for Rare Events with Applications*. Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.
- [7] Samuel Kotz, Tomasz J. Kozubowski, and Krzysztof Podgórski. *The Laplace Distribution and Generalizations*. Birkhäuser, Boston, 2001.

- [8] Tomasz J. Kozubowski and Krzysztof Podgórski. Asymmetric Laplace laws and modeling financial data. *Mathematical and Computer Modelling*, 34:1003–1021, 2001.
- [9] Tomasz J. Kozubowski and Svetlozar T. Rachev. The theory of geometric stable distributions and its use in modeling financial data. *European Journal of Operational Research*, 74:310–324, 1994.
- [10] Tomasz J. Kozubowski and Svetlozar T. Rachev. Multivariate geometric stable laws. *Journal of Computational Analysis and Applications*, 1(4):349–385, 1999.
- [11] Tomasz J. Kozubowski and Svetlozar T. Rachev. Univariate geometric stable laws. *Journal of Computational Analysis and Applications*, 1(2):177–217, 1999.
- [12] Jarl Waldemar Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15(1):211–225, 1922.
- [13] Stefan Mittnik and Svetlozar T. Rachev. Modeling asset returns with alternative stable distributions. *Econometric Review*, 12(3):261–330, 1993.
- [14] William J. Reed. The Pareto, Zipf and other power laws. *Economics Letters*, 74:15–19, 2001.
- [15] William J. Reed. The Pareto law of incomes—an explanation and an extension. *Physica A*, 319:469–486, 2003.
- [16] Domokos Szász. Limit theorems for the distributions of the sums of a random number of random variables. *Annals of Mathematical Statistics*, 43(6):1902–1913, 1972.
- [17] Alexis Akira Toda. Income dynamics with a stationary double Pareto distribution. *Physical Review E*, 83(4):046122, 2011.